

Imaging Informatics for Cell-Based Assays



B. Parvin

Imaging and Informatics
vision.lbl.gov

Lawrence Berkeley National Laboratory
Berkeley, CA 94720

Outline



- Motivation
 - Managing *information* (e.g., protein localization, physiological responses) from spatio-temporal data collected through optical microscopy
- Issues
 - How to acquire, annotate, compute, and organize *information* for meaningful representation
- Approach
 - Automated instrumentation, annotation, and on-line analysis (**VSOM**)
 - Integrated Informatics system for representing data, metadata, and quantitative information (**BioSig**)

- What proteins are being made in the cell (**identity**)? where are they expressed (**location**)? when are they active (**time**)? and what is their function (**activity**)?
 - Covalent modification
 - Kinetics of interactions
 - Monolayer (2D) or multicellular systems (3D)
- Optimization and screening of synthetic oligos for imaging mRNA
 - Specificity and hybridization efficiency
 - Signal to noise ratio
 - Pharmacokinetics (in and out fluxes)
- Optimization of microenvironment based on physiological responses
 - pH, temperature, reagents
 - Predictive models

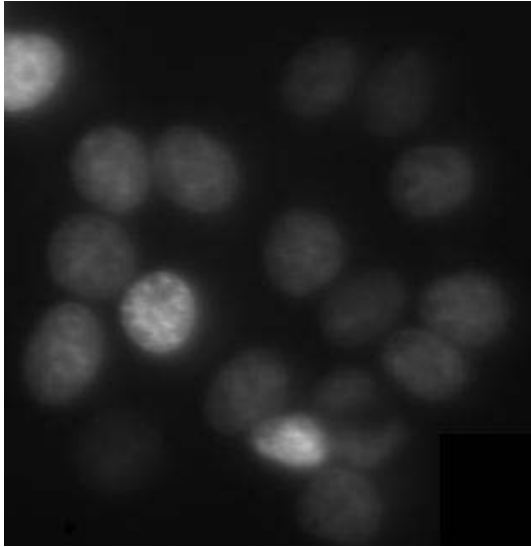
Challenges



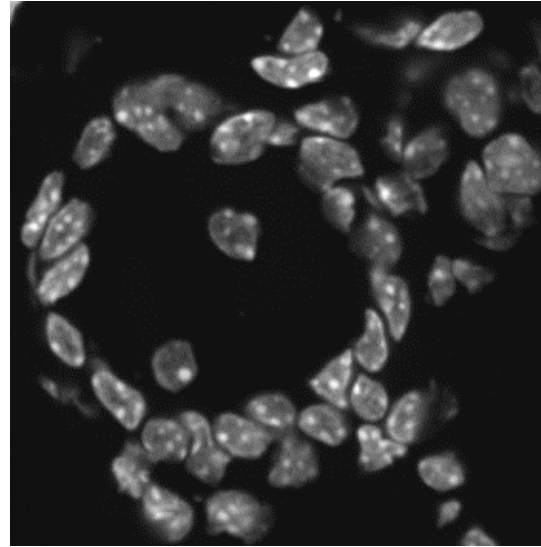
- Handling large volume of heterogeneous data
 - Standard molecular and cell biology techniques
 - Microscopy
 - Whole animal imaging
- Constructing and accessing complex schema and disparate ontologies
- Building visual routines for quantitative representation

❖ Biology is heterogeneous

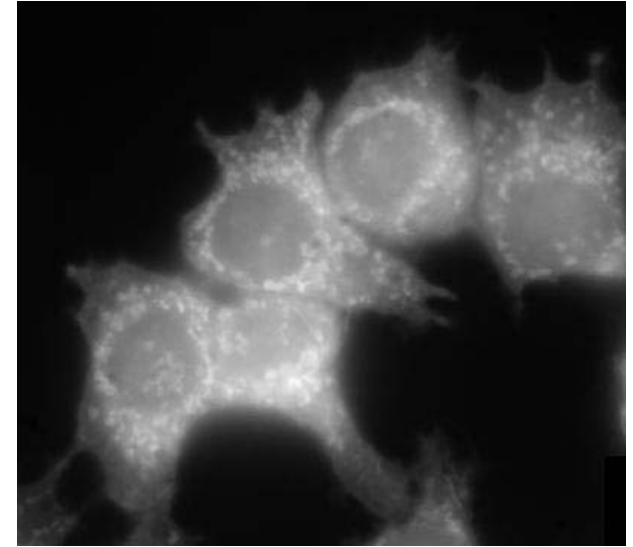
Challenges: quantitative representation for understanding biological images



Nuclear stain: image is noisy and the stain is not expressed uniformly



Nuclear stain: many internal substructures and overlapping compartments



Nuclear and mitochondrial: Image is noisy, which includes both random and speckled noise (internal substructures), and the stain is not expressed uniformly

Approach



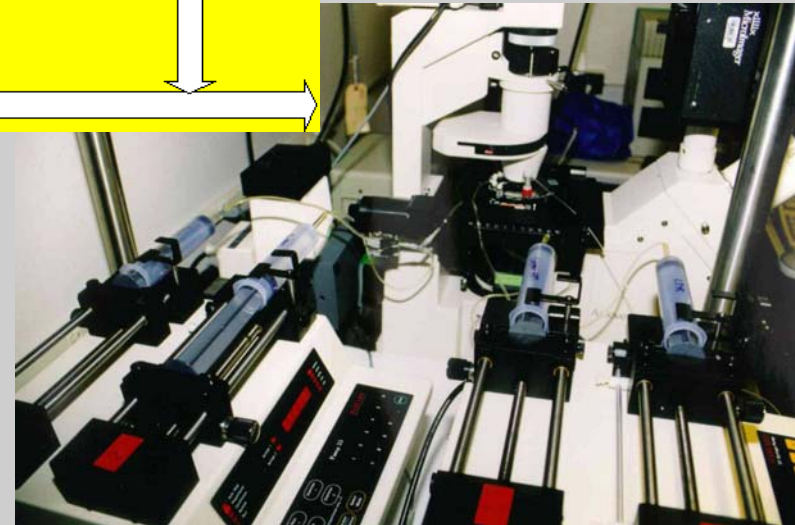
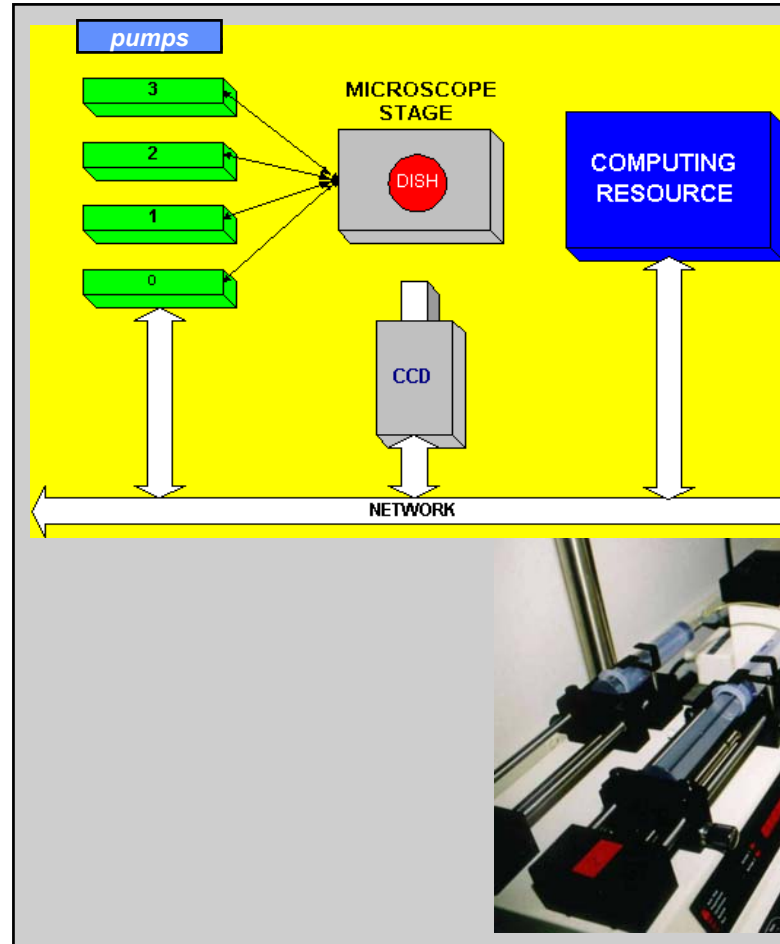
- Automated instrumentation for acquisition and control, on-line analysis, and standardized annotation
 - Close loop servo control
- Web-based architecture to access experimental annotations, data, and computed quantitative representation
 - Multilayer architecture
 - Schema and guided navigation through the database
 - Living cells and fixed tissue
- Novel algorithms to quantify features of interest using standardized interfaces
 - Quantitative visual routines for 5D datasets

Visual Servoing Optical Microscopy (VSOM)



Applications:

- Assay optimization
- In vitro screening



In vitro screening testbed

VSOM Interface



Vsom version 2.0

File Process Settings

Control Panel

Pos: ☐ 1 ☒ 2 ☒ 3 ☐ 4 ☐ 5 ☐ 6 ☒ X

Wavelength: N/A N/A N/A N/A N/A N/A N/A

Exposure time:

Sampling rate: Duration:

Diagnostics

Pos	1	2	3	4	5	6
						N/A
						N/A
						N/A

Process step: 3

ICC Imaging System

File Process

Control Panel

Camera: ☐ CCD ☒ Cooled CCD

Filter: ☒ 1 ☒ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☒ Trans

Exposure time:

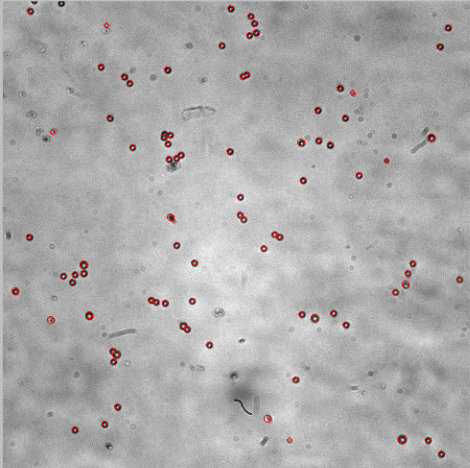
Trans:

Display:

Capture! Save

Annotation

Calibration with beads and measuring leakage between different frequencies



Calibration Step: 1

Configure VSOM

General information Instrument Filter VSOM schedule Specimen/Vessel

Reagent Group

Syringes	Diameter	Volume
2	26.7	60

Add syringe Undo

Syringe components

Fluorophore	Component	Concentration	Unit
Yopro1	SLO	20	ng

Components: 1 ☒ Add Component Undo

Schedule

Intervals	Description	Syringe	Start time	Duration	Rate	Unit	Washout
1	buffer	1	10	600	1	MV	T

Add schedule Undo

Global trigger

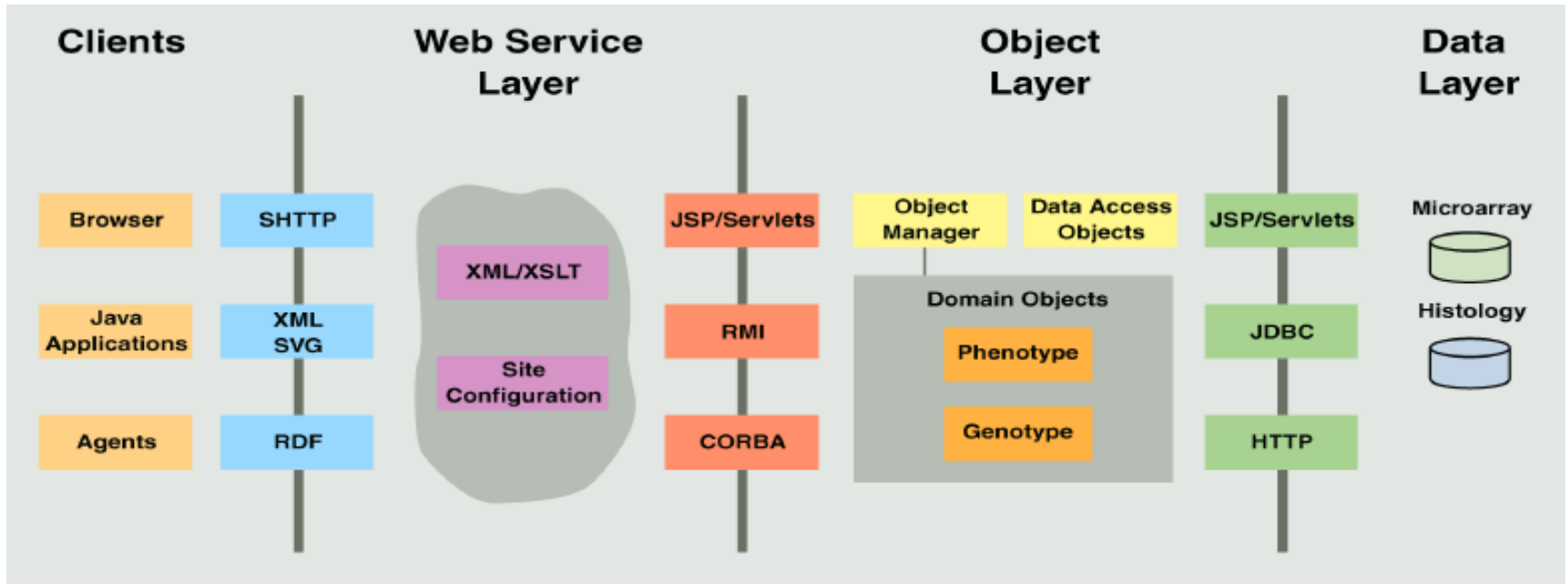
If Population: 50 Unit: bot Pos: 3 Condition: >= Value: 40 or Max duration: Interval: 1 then Triggers: 1

Add trigger Undo

OK Cancel

Annotation Step: 2

Static and functional structure



Functional structure

Data/semantics standards

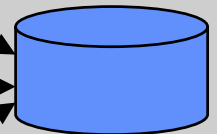
Standard IO format

Graphic-based

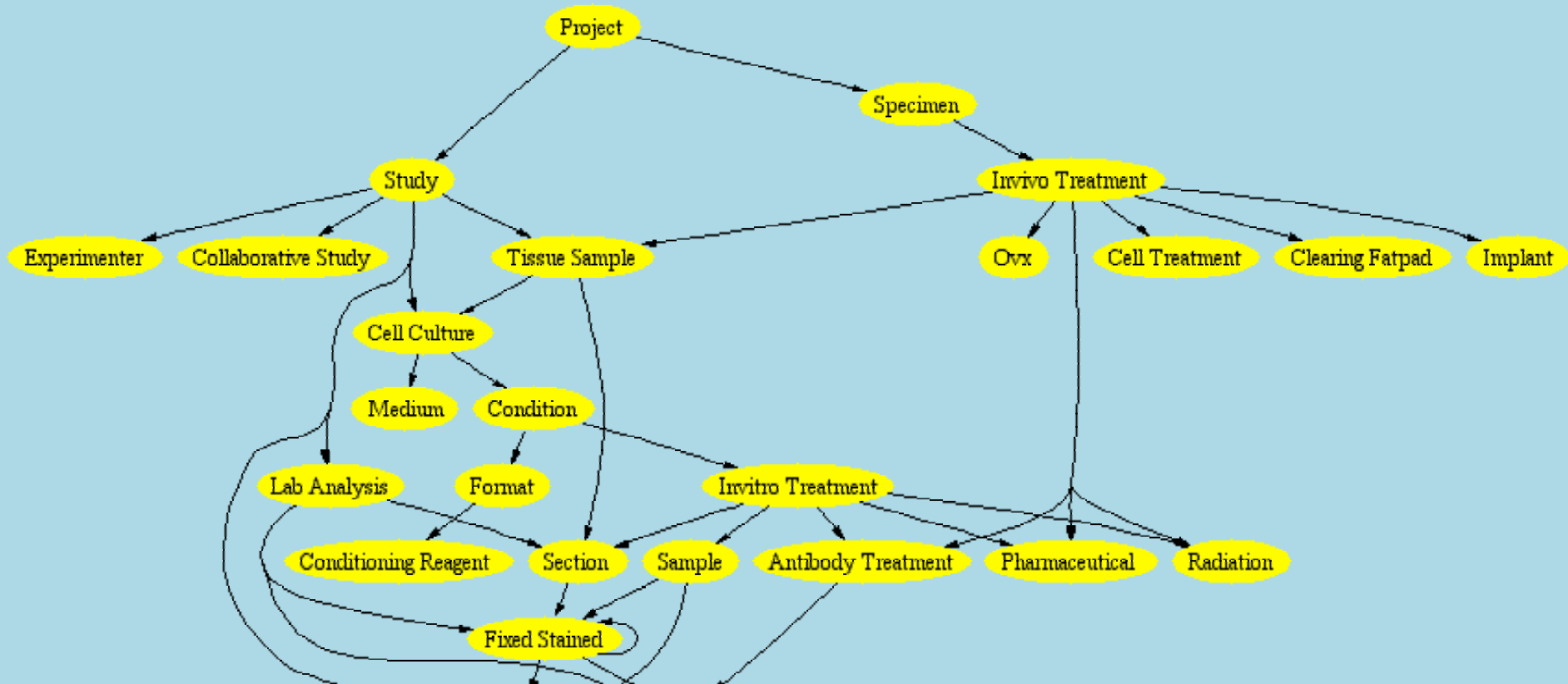
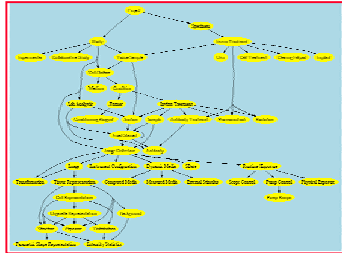
Data acquisition and annotation

Data analysis and representation

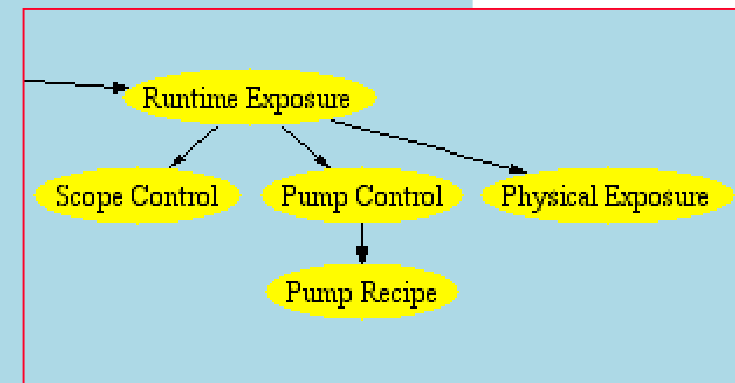
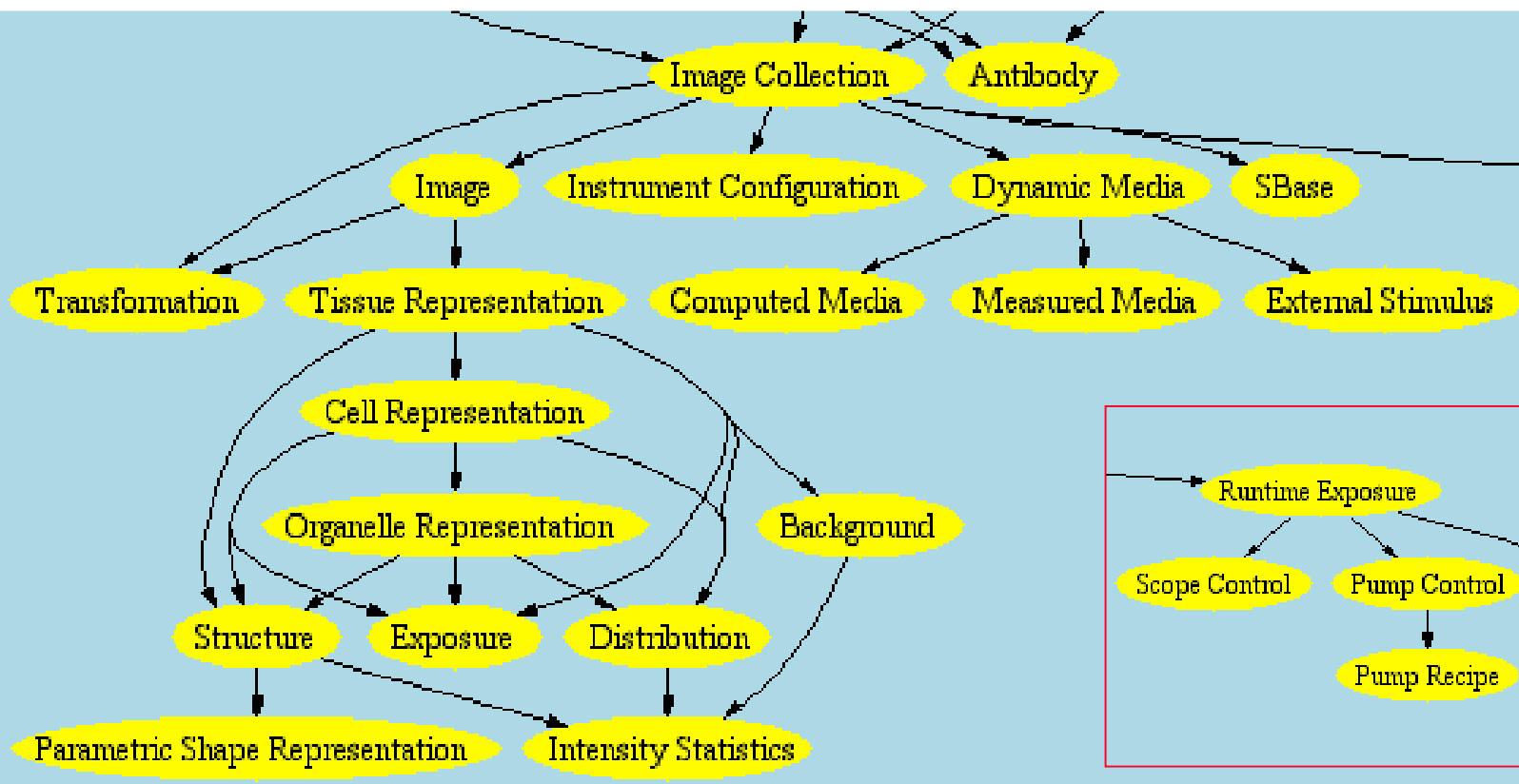
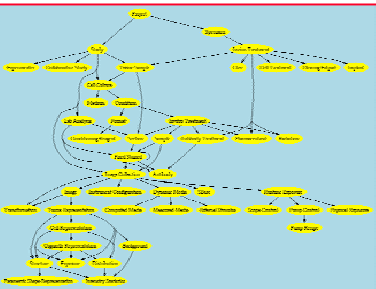
Visualization and query



Schema



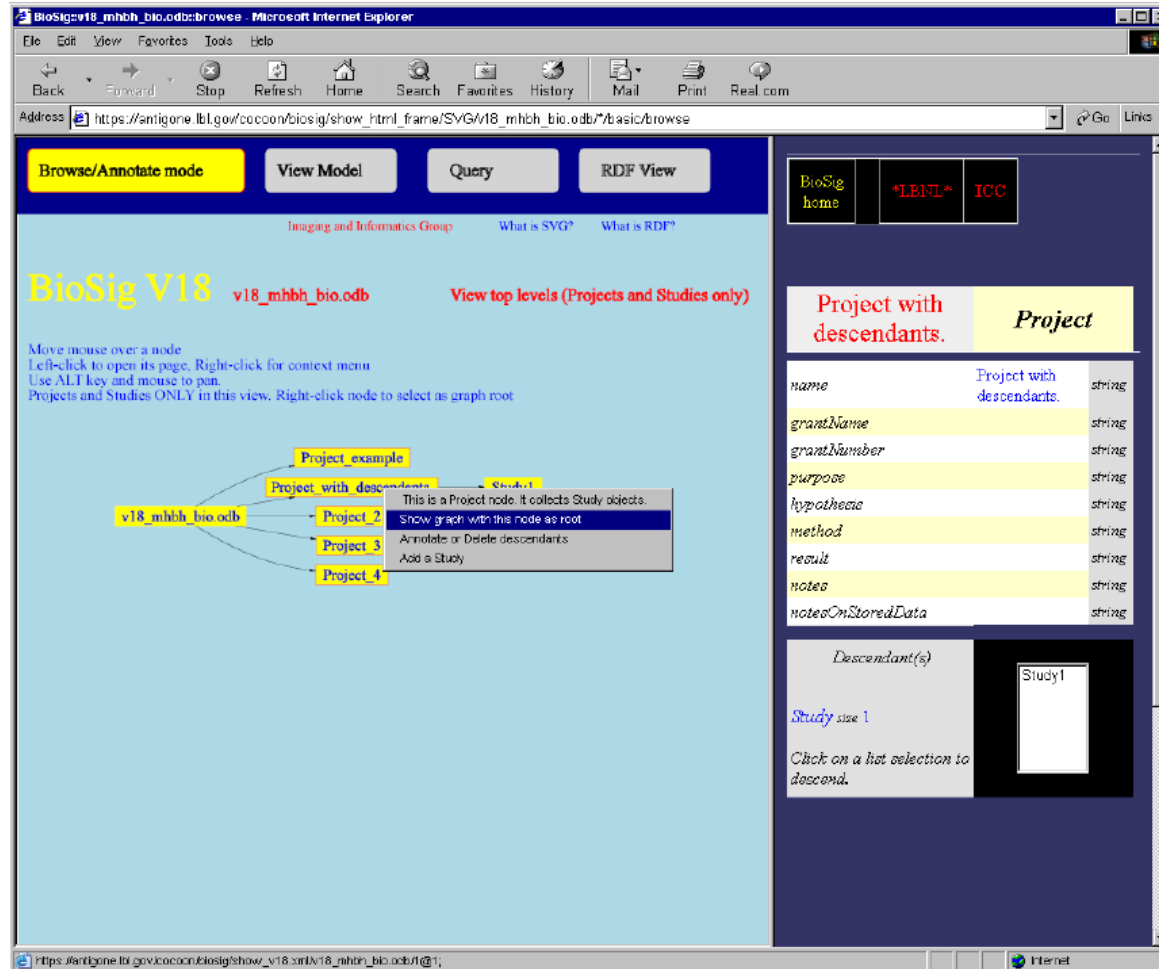
Schema



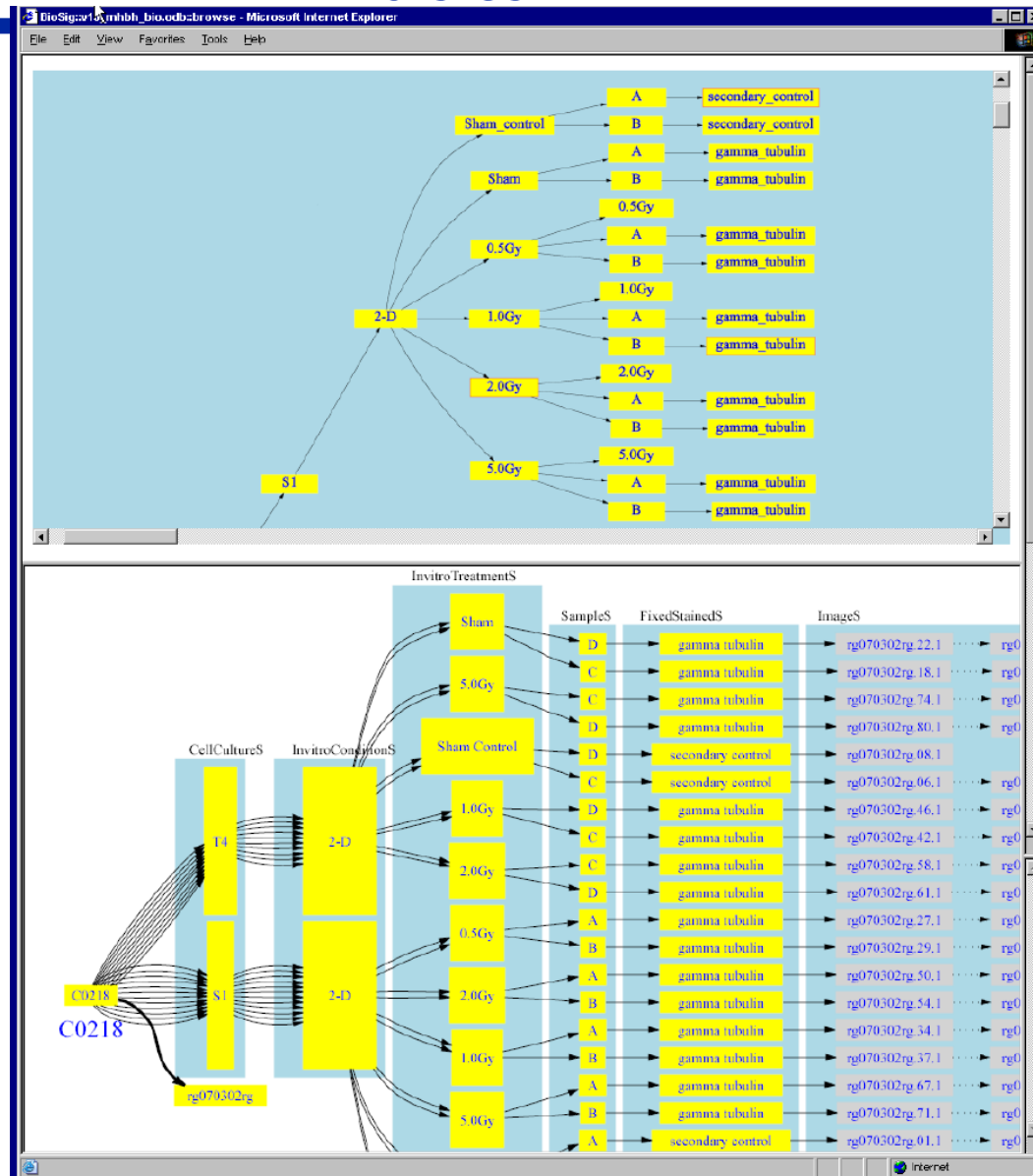
Dynamic user interface for annotating experimental images



- GUI feature
 - Declarative
 - Graphics
- Database
 - Postgress
 - Objectstore
- Data standards/ontologies
 - OME
 - NLM

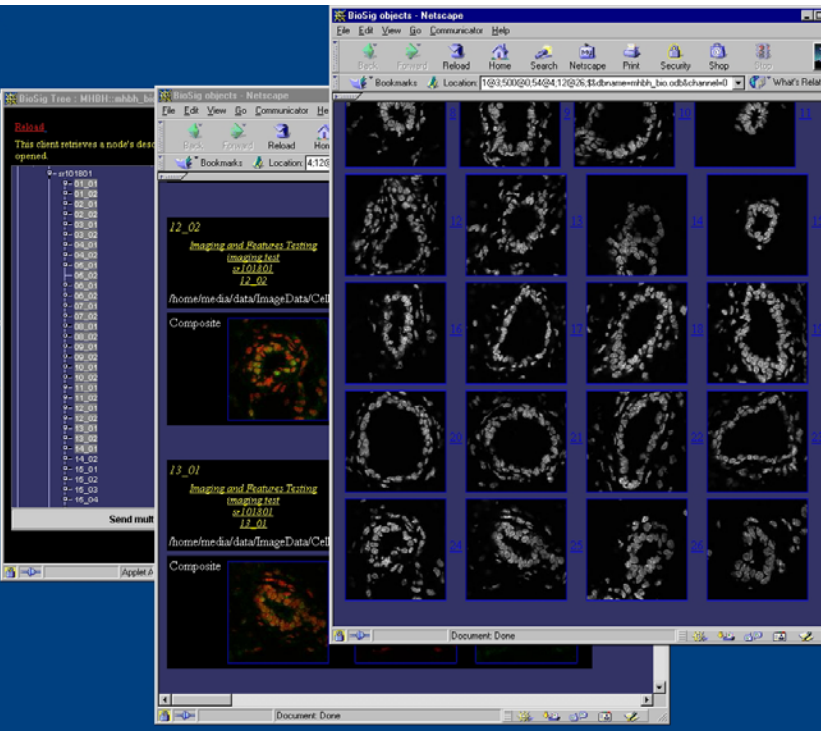
A screenshot of the BioSig V18 web application running in a Microsoft Internet Explorer browser. The browser address bar shows the URL: https://antigone.lbl.gov/cocoon/biosig/show_html_frame/SVG/v18_mhbbh_bio.odb/?basic/browse. The application interface has a dark blue header with buttons for "Browse/Annotate mode", "View Model", "Query", and "RDF View". Below the header, there's a light blue area with the text "BioSig V18 v18_mhbbh_bio.odb" and "View top levels (Projects and Studies only)". A diagram shows a hierarchical structure with "v18_mhbbh_bio.odb" at the top, branching into "Project example", "Project with descendants", "Project_2", "Project_3", and "Project_4". A context menu is open over "Project_2", showing options: "This is a Project node. It collects Study objects.", "Show graph with this node as root", "Annotate or Delete descendants", and "Add a Study". On the right side, there's a sidebar with "BioSig home", "*LENT*", and "ICC" buttons. Below these, there's a section titled "Project with descendants." and "Project", followed by a table with columns "name", "Project with descendants.", and "string". The table lists various project attributes like "grantName", "grantNumber", "purpose", "hypothesis", "method", "result", "notes", and "notesOnStoredData". At the bottom of the sidebar, there's a section titled "Descendant(s)" and "Study1", with a button labeled "Click on a list selection to descend."

Navigation through experimental data

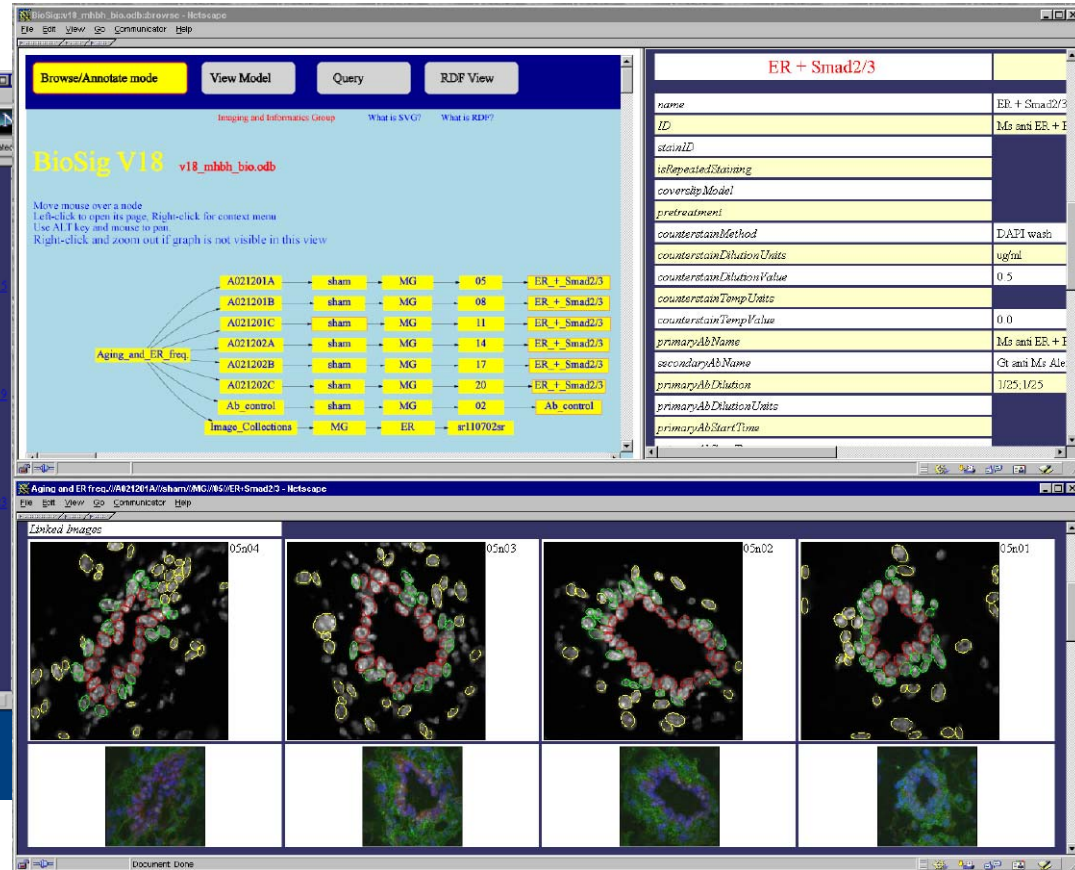


Feb. 5th, 2004

Presentation layer

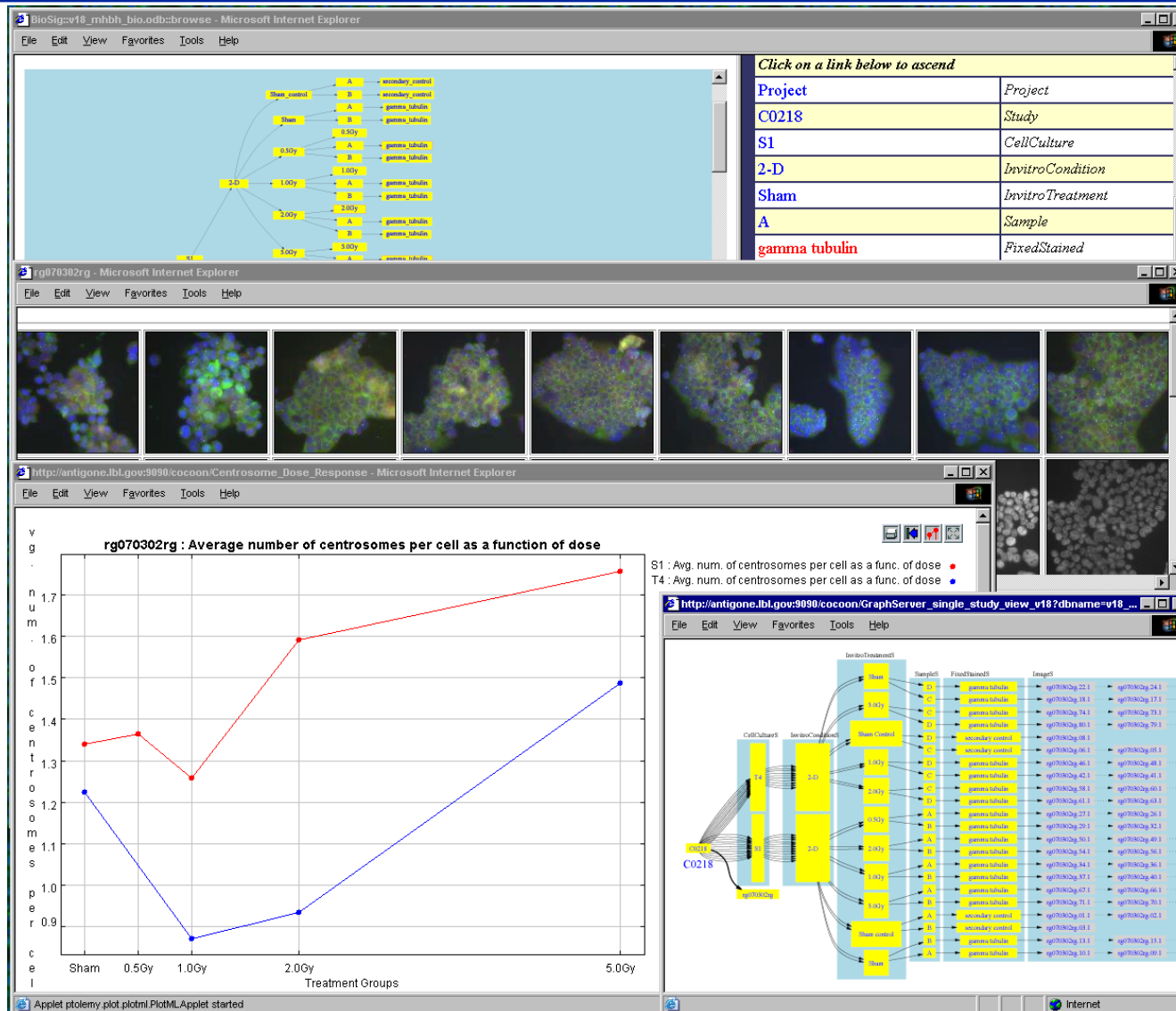


Automated scaling



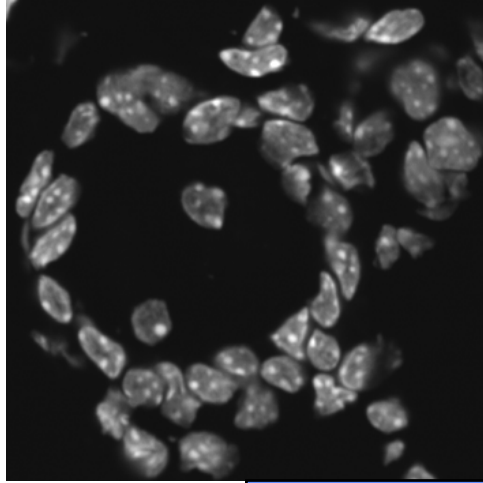
Raw and processed data

Integrated analysis and browsing of data

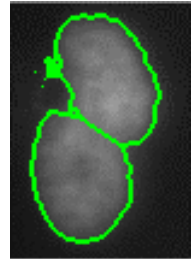
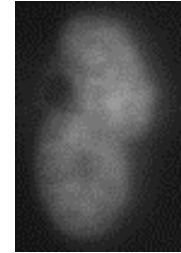


Feb. 3rd, 2004

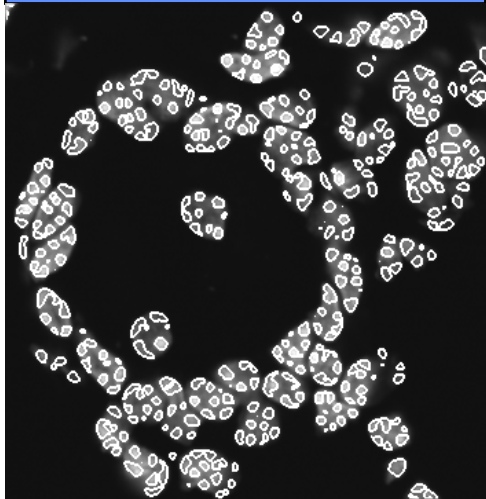
Steps in quantitative representation (geometry-based –1)



Original image

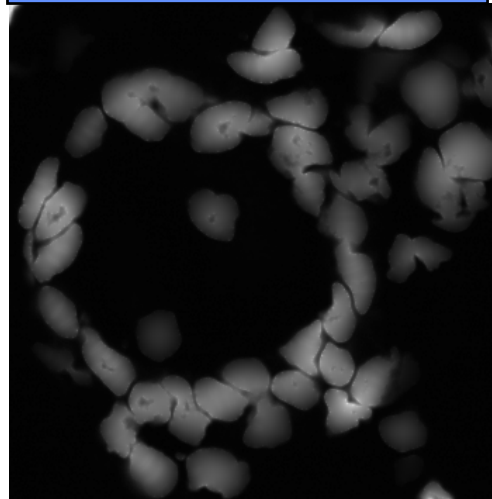


**EXTRACT ELLIPTIC
REGIONS**



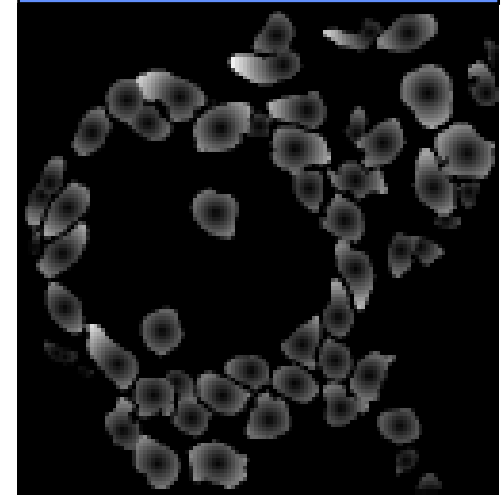
Detect small substructures

HARMONIC CUTS



Interpolate substructures

CENTROID TRANSFORM

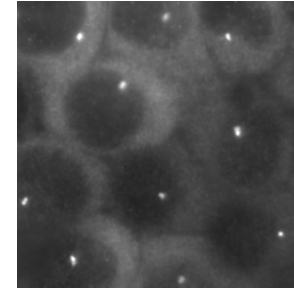


Cluster by geometry

Steps in quantitative representation (statistical learning)

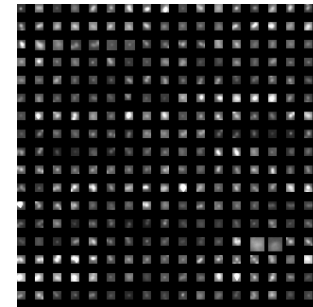
1.

Nuclear localization



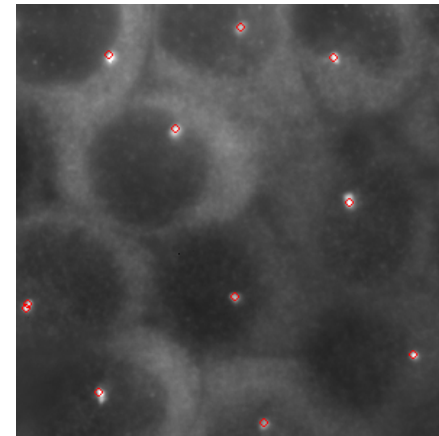
2.

*Construct training
database punctate
events*



3.

*Use context to assign
punctate events to each
cell*



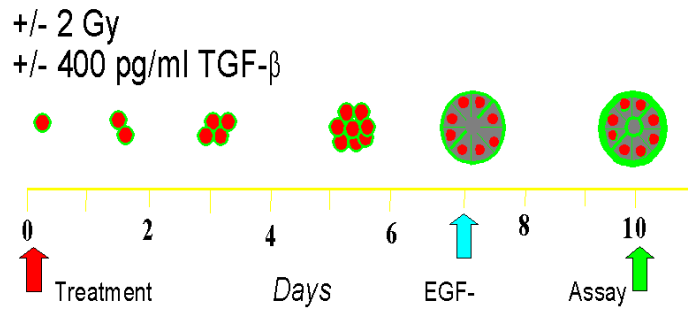
Applications



- Quantitative representation of 2D and 3D data
 - Cell-cell communication in multicellular systems
- Time-resolved (4D data) responses in living cells
 - Physiological responses
 - mRNA imaging

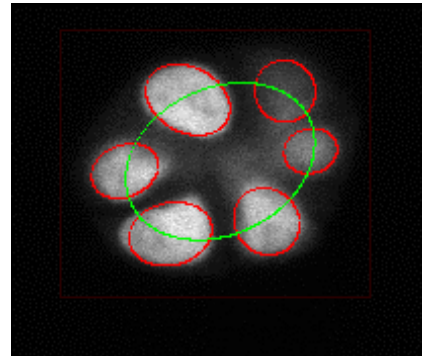
Characterizing colony formation

Experimental Protocol

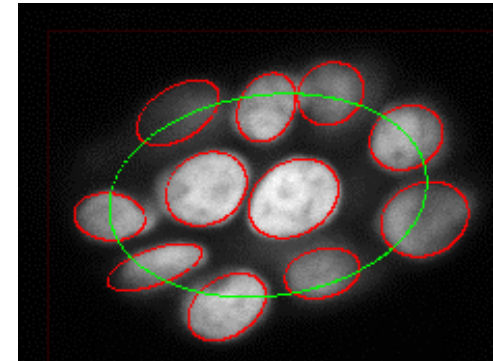


**Ellipsoidal
representation**

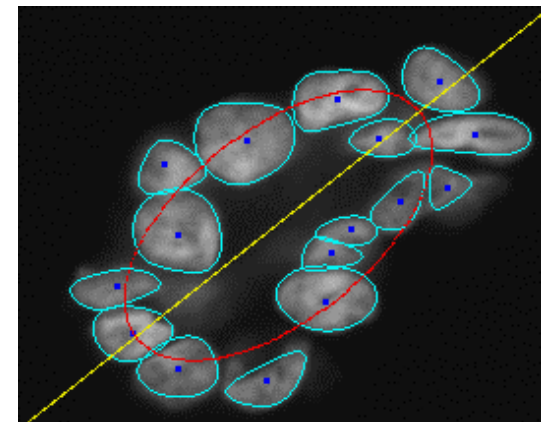
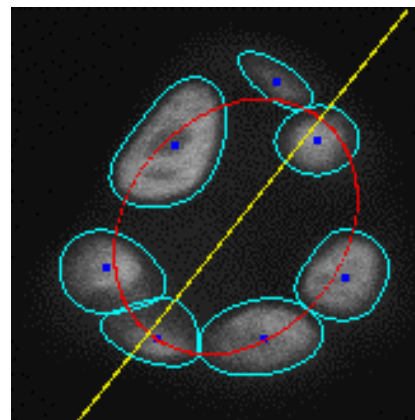
***Normal organization
of a colony***



***Altered organization
of colony***

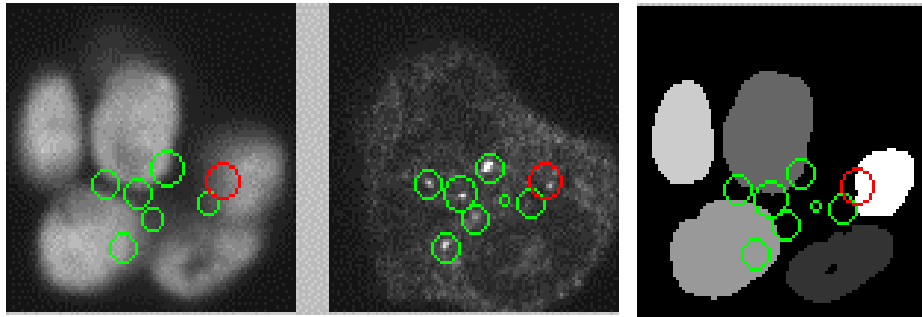


**Hyperquadric
representation**

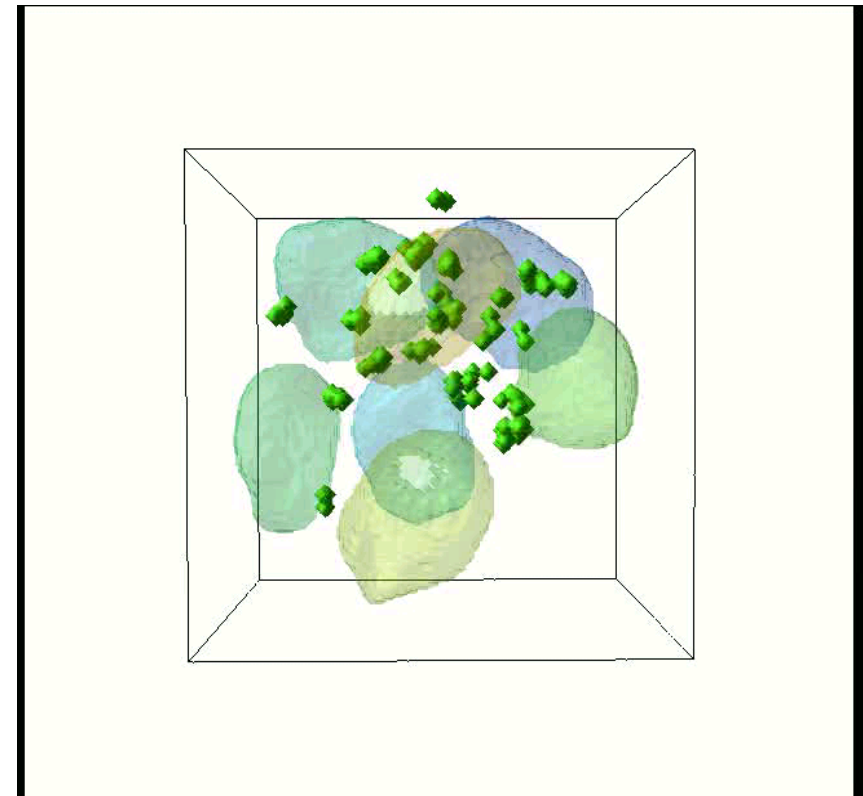


Characterizing cell-cell communication

- Collect training samples for connexin expression
- Use statistics, shape features, and context to detect connexin

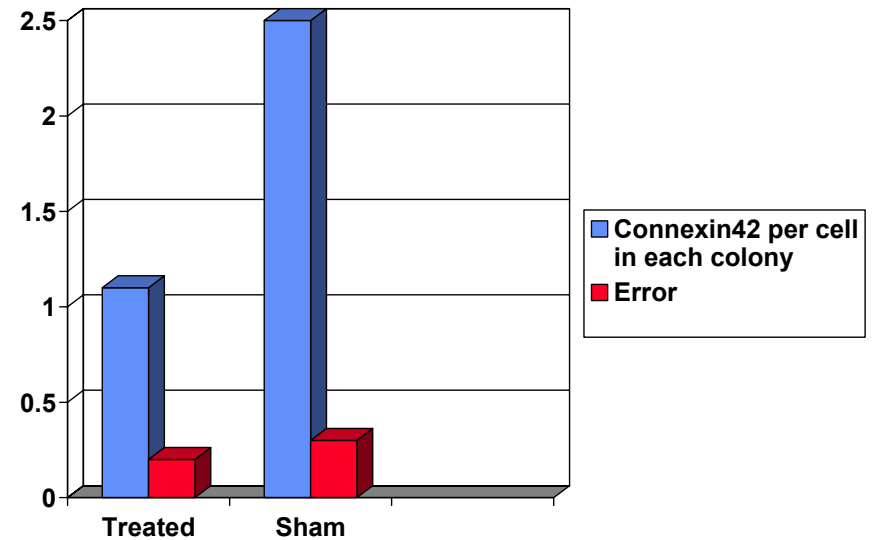
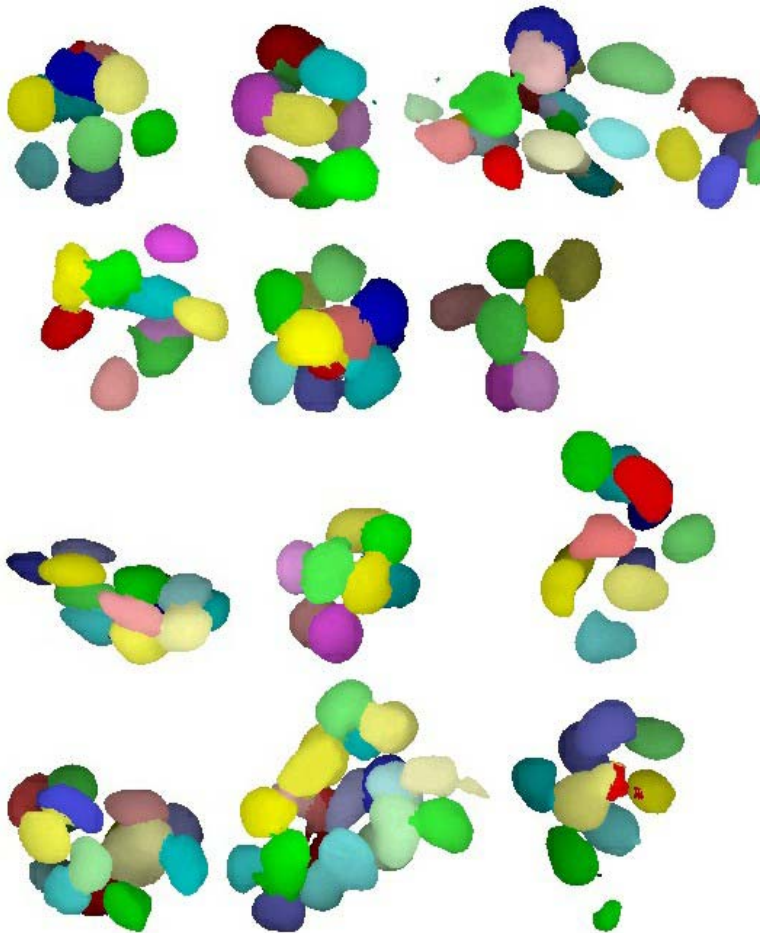


One slice of 3d volume at 360 and 480 nm



Computed 3D structure of cultured cells and protein localization

Population studies

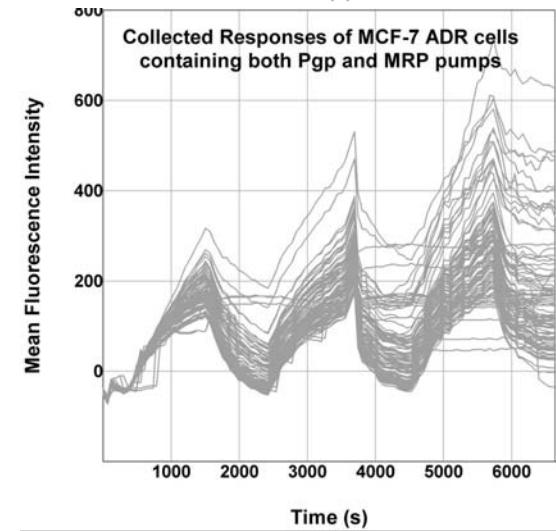
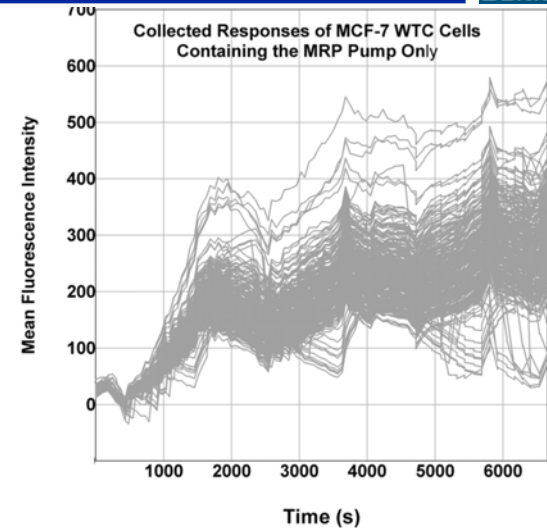
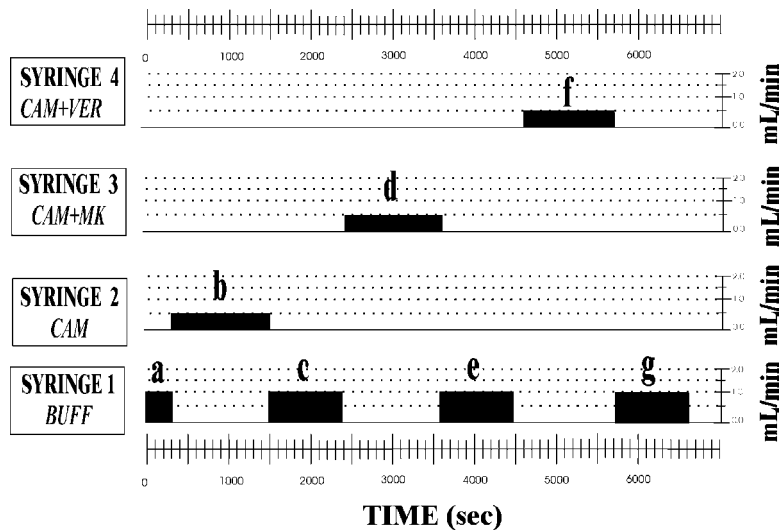
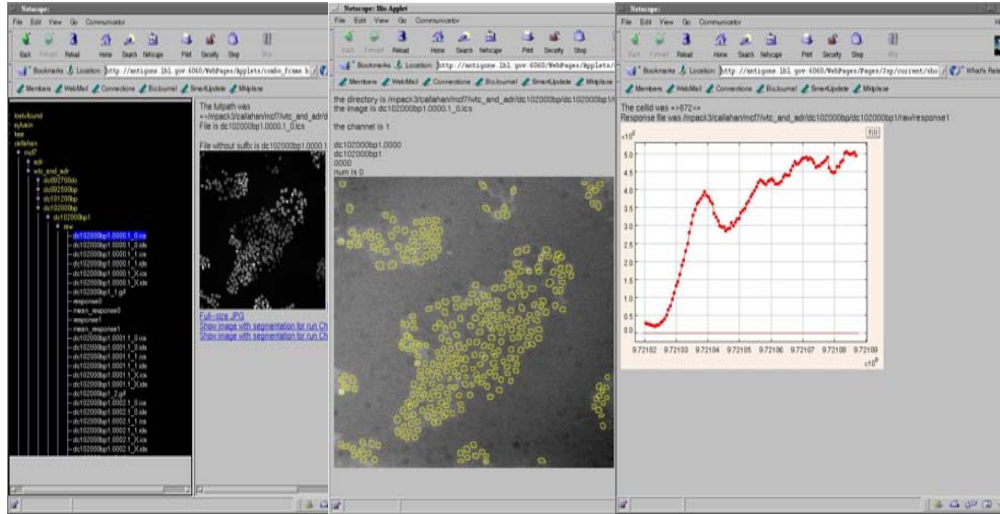


Population study indicates that number of gap junctions are reduced as a function of particular treatment (one experiment with ~4 Gbytes of data)

Characterizing physiological fingerprints (CAL-CAM assay)



Kinetic uptake analysis for a single live cell



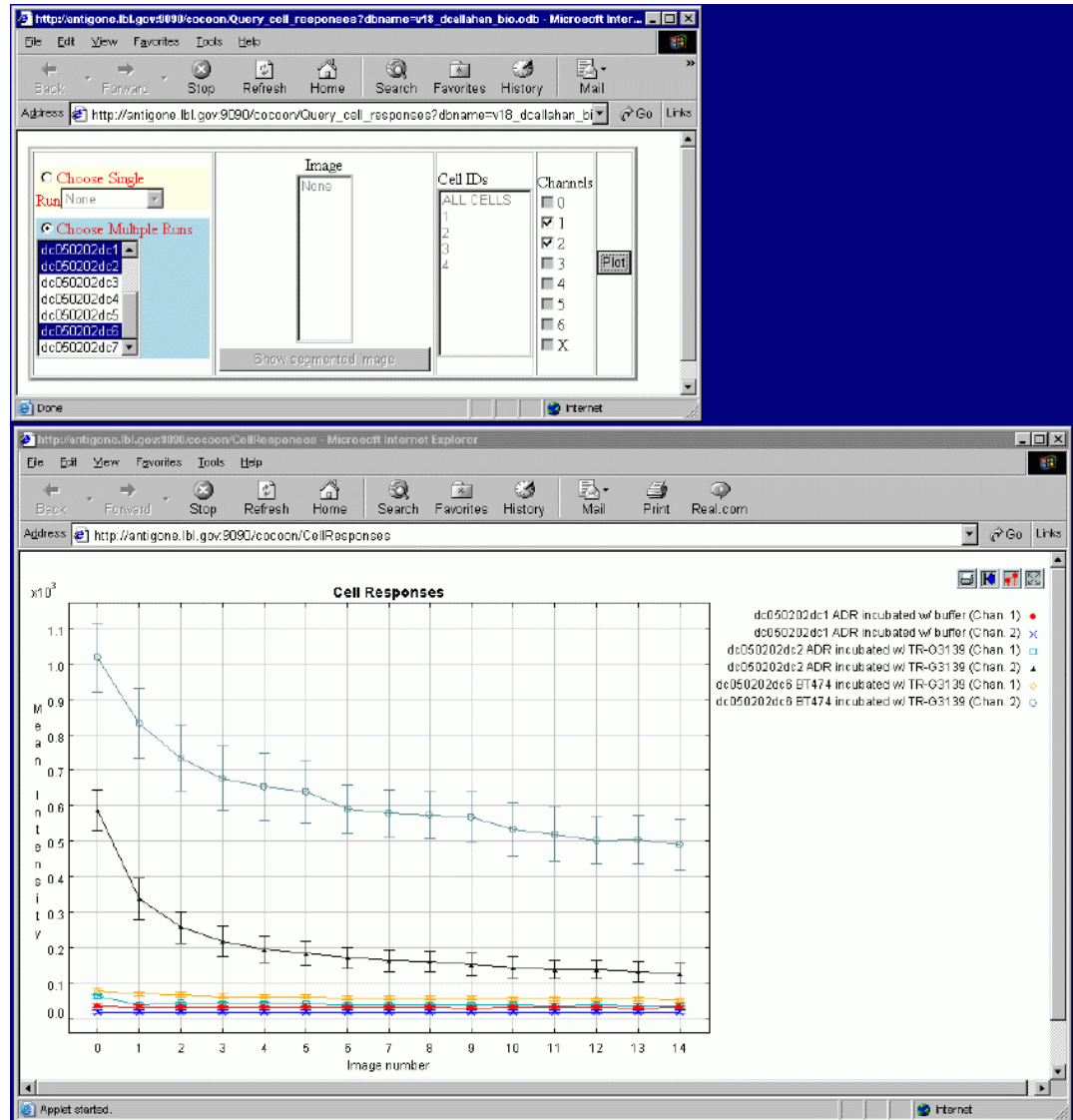
Kinetic uptakes for two different cell lines

Query operators for comparative analysis

Plots indicate washout curves for three different cell lines

Size of data being compared:

~400 Mbytes



Summary/Acknowledgement



- Cell-based assays have the potential to generate large volume of complex heterogeneous data
 - Automated instrumentation and standard ontologies
 - Novel quantitative methods for representation
 - Informatics infrastructure
- Developers
 - G. Fontenay
 - R. Romano
 - Q. Yang
 - B. Parvin
- Collaborators
 - D. Callahan (LBNL)
 - M.H. Barcellos-Hoff (LBNL)
 - D. Sudar (LBNL)
 - M. Nielson-Hamilton (AMES)
 - A. Gifford (BNL)
 - T. Ilstiy (UCSF)
- Credits:
 - Project is funded by Department of Energy Office of Biological and Environment Research
 - Life/Medical Sciences Division